Testing the Equality of Covariance Operators in Functional Samples*

STEFAN FREMDT

Mathematical Institute, University of Cologne

LAJOS HORVÁTH

Department of Mathematics, University of Utah

PIOTR KOKOSZKA

Department of Mathematics and Statistics, Utah State University

JOSEF G. STEINEBACH

Mathematical Institute, University of Cologne

ABSTRACT. **We propose a robust test for the equality of the covariance structures in two functional samples. The test statistic has a chi-square asymptotic distribution with a known number of degrees of freedom, which depends on the level of dimension reduction needed to represent the data. Detailed analysis of the asymptotic properties is developed. Finite sample performance is examined by a simulation study and an application to egg–laying curves of fruit flies.**

Keywords: Asymptotic distribution, Covariance operator, Functional data, Quadratic forms, Two sample problem.

Abbreviated Title: Equality of covariance operators

AMS subject classification: Primary 62G10; secondary 62G20, 62H15

1. Introduction

The last decade has seen increasing interest in methods of functional data analysis which offer novel and effective tools for dealing with problems where curves can naturally be viewed as data objects. The books by Ramsay & Silverman (2005) and Ramsay et al. (2009) offer comprehensive introductions to the subject, the collection Ferraty & Romain (2011) reviews some recent developments focusing on advances in the relevant theory, while the monographs of Bosq (2000), Ferraty & Vieu (2006) and Horváth & Kokoszka (2011+) develop the field in several important directions. Despite the emergence of many alternative ways of looking at functional data, and many dimension reduction approaches, the functional principal components (FPC's) still remain the most important starting

point for many functional data analysis procedures, Reiss & Ogden (2007), Gervini (2008), Yao & Müller (2010), Gabrys et al. (2010) are just a handful of illustrative references. The FPC's are the eigenfunctions of the covariance operator. This paper focuses on testing if the covariance operators of two functional samples are equal. By the Karhunen-Loève expansion, this is equivalent to testing if both samples have the same set of FPC's. Benko et al. (2009) developed bootstrap procedures for testing the equality of specific FPC's. Panaretos et al. (2010) proposed a test of the type we consider, but assuming that the curves have a Gaussian distribution. The main result of Panaretos et al. (2010) follows as a corollary of our more general approach (Theorem 2).

Despite their importance, two sample problems for functional data received relatively little attention. In addition to the work of Benko et al. (2009) and Panaretos et al. (2010), the relevant references are Horváth et al. (2009) and Horváth et al. (2011) who focus, respectively, on the regression kernels in functional linear models and the mean of functional data exhibiting temporal dependence. Clearly, if some population parameters of two functional samples are different, estimating them using the pooled sample may lead to spurious conclusions. Due to the importance of the FPC's, a relatively simple and robust procedure for testing the equality of the covariance operators is called for.

The remainder of this paper is organized as follows. Section 2. sets out the notation and definitions. The construction of the test statistic and its asymptotic properties are developed in Section 3.. Section 4. reports the results of a simulation study and illustrates the procedure by application to egg-laying curves of Mediterranean fruit flies. The proofs of the asymptotic results of Section 3. are given in Section 5..

2. Preliminaries

Let $X_1, X_2, \ldots, X_N$ be independent, identically distributed random variables in $L_2[0,1]$ with $EX_i(t) = \mu(t)$ and $cov(X_i(t), X_i(s)) = C(t, s)$. We assume that another sample $X_1^*, X_2^*, \ldots X_M^*$ is also available and let $\mu^*(t) = EX_i^*(t)$ and $C^*(t, s) = cov(X_i^*(t), X_i^*(s))$ for $t, s \in [0, 1]$. We wish to test the null hypothesis

$$H_0: \ C = C^*$$

against the alternative H_A that H_0 does not hold.

A crucial assumption considering the asymptotics of our test procedure will be that

$$\Theta_{N,M} = \frac{N}{M + N} \to \Theta \in (0, 1) \quad \text{as} \quad N, M \to \infty. \tag{1}$$

For the construction of our test procedure we will use an estimate of the asymptotic pooled covariance operator $\mathfrak{R}$ of the two given samples (cf. (4)) which is defined by the kernel

$$R(t, s) = \Theta C(t, s) + (1 - \Theta)C^*(t, s).$$

Denote by $(\lambda_1, \varphi_1), (\lambda_2, \varphi_2), \ldots$ the eigenvalue/eigenfunction pairs of $\mathfrak{R}$, which are defined by

$$\lambda_k \varphi_k(t) = \mathfrak{R}\varphi_k(t) = \int_0^1 R(t, s)\varphi_k(s)ds, \quad t \in [0, 1], \quad 1 \le k < \infty . \tag{2}$$

Throughout this paper we assume

$$\lambda_1 > \lambda_2 > \ldots > \lambda_p > \lambda_{p+1}, \tag{3}$$

i.e. there exist at least p distinct (positive) eigenvalues. Under assumption (3), we can uniquely (up to signs) choose $\varphi_1, \ldots, \varphi_p$ satisfying (2), if we require $\|\varphi_i\| = 1$, where for a positive integer d and for $x \in L_2\left([0,1]^d\right)$

$$\|x\| = \left(\int_0^1 \cdots \int_0^1 x^2(t_1, \ldots, t_d)dt_1 \cdots dt_d\right)^{1/2}.$$

Thus, under (3), $\{\varphi_i, 1 \leq i \leq p\}$ is an orthonormal system that can be extended to an orthonormal basis $\{\varphi_i, 1 \leq i < \infty\}$.

If H_0 holds, then (λ_i, φ_i), $1 \leq i < \infty$, are also the eigenvalues/eigenfunctions of the covariance operators $\mathfrak{C}$ of the first and $\mathfrak{C}^*$ of the second sample. To construct a test statistic which converges under H_0, we can therefore pool the two samples, as explained in Section 3..

3. The test and the asymptotic results

Our procedure is based on projecting the observations onto a suitably chosen finite-dimensional space. To define this space, introduce the empirical pooled covariance operator $\widehat{\mathfrak{R}}_{N,M}$ defined by the kernel

$$\widehat{R}_{N,M}(t,s) = \frac{1}{N+M}\left\{ \sum_{k=1}^N (X_k(t) - \overline{X}_N(t))(X_k(s) - \overline{X}_N(s)) \right. \tag{4}$$
$$\left. + \sum_{k=1}^M (X_k^*(t) - \overline{X}_M^*(t))(X_k^*(s) - \overline{X}_M^*(s)) \right\},$$

where

$$\overline{X}_N(t) = \frac{1}{N}\sum_{k=1}^N X_k(t) \quad \text{and} \quad \overline{X}_M^*(t) = \frac{1}{M}\sum_{k=1}^M X_k^*(t)$$

are the sample mean functions. Let $(\widehat{\lambda}_i, \widehat{\varphi}_i)$ denote the eigenvalues/eigenfunctions of $\widehat{\mathfrak{R}}_{N,M}$, i.e.

$$\widehat{\lambda}_i \widehat{\varphi}_i(t) = \widehat{\mathfrak{R}}_{N,M} \widehat{\varphi}_i(t) = \int_0^1 \widehat{R}_{N,M}(t,s)\widehat{\varphi}_i(s)ds, \quad t \in [0,1], \quad 1 \leq i \leq N+M,$$

with $\widehat{\lambda}_1 \geq \widehat{\lambda}_2 \geq \ldots$. We can and will assume that the $\widehat{\varphi}_i$ form an orthonormal system. We consider the projections

$$\widehat{a}_k(i) = \langle X_k - \overline{X}_N, \widehat{\varphi}_i \rangle = \int_0^1 (X_k(t) - \overline{X}_N(t))\widehat{\varphi}_i(t)dt \tag{5}$$

and

$$\widehat{a}_k^*(j) = \langle X_k^* - \overline{X}_M^*, \widehat{\varphi}_j \rangle = \int_0^1 \left(X_k^*(t) - \overline{X}_M^*(t) \right) \widehat{\varphi}_j(t) dt. \tag{6}$$

To test H_0, we compare the matrices $\widehat{\Delta}_N$ and $\widehat{\Delta}_M^*$ with entries

$$\widehat{\Delta}_N(i,j) = \frac{1}{N} \sum_{k=1}^N \widehat{a}_k(i)\widehat{a}_k(j), \quad 1 \le i, j \le p,$$

and

$$\widehat{\Delta}_M^*(i,j) = \frac{1}{M} \sum_{k=1}^M \widehat{a}_k^*(i)\widehat{a}_k^*(j), \quad 1 \le i, j \le p.$$

We note that $\widehat{\Delta}_N(i,j) - \widehat{\Delta}_M^*(i,j)$ is the projection of $\widehat{C}_N(t,s) - \widehat{C}_M^*(t,s)$ in the direction of $\widehat{\varphi}_i(t)\widehat{\varphi}_j(s)$, where

$$\widehat{C}_N(t,s) = \frac{1}{N} \sum_{k=1}^N \left(X_k(t) - \overline{X}_N(t) \right) \left(X_k(s) - \overline{X}_N(s) \right)$$

and

$$\widehat{C}_M^*(t,s) = \frac{1}{M} \sum_{k=1}^M \left(X_k^*(t) - \overline{X}_M^*(t) \right) \left(X_k^*(s) - \overline{X}_M^*(s) \right)$$

are the empirical covariances of the two samples.

From the columns below the diagonal of $\widehat{\Delta}_N - \widehat{\Delta}_M^*$ we create a vector $\widehat{\boldsymbol{\xi}}_{N,M}$ as follows:

$$\widehat{\boldsymbol{\xi}}_{N,M} = \text{vech}\left(\widehat{\Delta}_N - \widehat{\Delta}_M^* \right) = \begin{pmatrix} \widehat{\Delta}_N(1,1) - \widehat{\Delta}_M^*(1,1) \\ \widehat{\Delta}_N(2,1) - \widehat{\Delta}_M^*(2,1) \\ \vdots \\ \widehat{\Delta}_N(p,p) - \widehat{\Delta}_M^*(p,p) \end{pmatrix}. \tag{7}$$

For the properties of the vech operator we refer to Abadir & Magnus (2005).

Next we estimate the asymptotic covariance matrix of $(MN/(N+M))^{1/2}\widehat{\boldsymbol{\xi}}_{N,M}$. Let

$$\widehat{L}_{N,M}(k,k') = (1 - \Theta_{N,M}) \left\{ \frac{1}{N} \sum_{\ell=1}^N \widehat{a}_\ell(i)\widehat{a}_\ell(j)\widehat{a}_\ell(i')\widehat{a}_\ell(j') - \langle \widehat{\mathfrak{C}}_N \widehat{\varphi}_i, \widehat{\varphi}_j \rangle \langle \widehat{\mathfrak{C}}_N \widehat{\varphi}_{i'}, \widehat{\varphi}_{j'} \rangle \right\}$$

$$+ \Theta_{N,M} \left\{ \frac{1}{M} \sum_{\ell=1}^M \widehat{a}_\ell^*(i)\widehat{a}_\ell^*(j)\widehat{a}_\ell^*(i')\widehat{a}_\ell^*(j') - \langle \widehat{\mathfrak{C}}_M^* \widehat{\varphi}_i, \widehat{\varphi}_j \rangle \langle \widehat{\mathfrak{C}}_M^* \widehat{\varphi}_{i'}, \widehat{\varphi}_{j'} \rangle \right\}$$

where i, j, i', j' depend on k, k' (see below), and $\widehat{\mathfrak{C}}_N$ $(\widehat{\mathfrak{C}}_M^*)$ is interpreted as an operator with $\widehat{\mathfrak{C}}_N$ defined as

$$\widehat{\mathfrak{C}}_N \widehat{\varphi}_i = \int_0^1 \widehat{C}_N(t,s)\widehat{\varphi}_i(s) ds.$$

(An analogous definition holds for $\widehat{\mathfrak{C}}_M^*$.) From this definition it follows that

$$\langle \widehat{\mathfrak{C}}_N \widehat{\varphi}_i, \widehat{\varphi}_j \rangle = \frac{1}{N} \sum_{l=1}^{N} \widehat{a}_\ell(i) \widehat{a}_\ell(j).$$

We note that one can use $\widehat{L}_{N,M}^*(k, k')$ instead of $\widehat{L}_{N,M}(k, k')$, where $\widehat{L}_{N,M}^*(k, k')$ is defined like $\widehat{L}_{N,M}(k, k')$, but $\langle \widehat{\mathfrak{C}}_N \widehat{\varphi}_i, \widehat{\varphi}_j \rangle$ and $\langle \widehat{\mathfrak{C}}_M^* \widehat{\varphi}_i, \widehat{\varphi}_j \rangle$ are replaced with 0 if $i \neq j$ and $\widehat{\lambda}_i$ if $i = j$. In the same spirit, $\langle \widehat{\mathfrak{C}}_N \widehat{\varphi}_{i'}, \widehat{\varphi}_{j'} \rangle$ and $\langle \widehat{\mathfrak{C}}_M^* \widehat{\varphi}_{i'}, \widehat{\varphi}_{j'} \rangle$ are replaced with 0 for $i' \neq j'$ and $\widehat{\lambda}_{i'}$ if $i' = j'$.
The index (i, j) is computed from k in the following way: Let

$$k' = \frac{p(p+1)}{2} - k + 1, \quad i' = p - i + 1, \quad \text{and} \quad j' = p - j + 1. \tag{8}$$

We look at an upper triangle matrix $(a_{i',j'})$. Then, for column j', we have that $(j' - 1)j'/2 < k \leq j'(j' + 1)/2$. Thus $j' = \left\lceil \sqrt{2k' + \frac{1}{4}} - \frac{1}{2} \right\rceil$ and $i' = k' - (j' - 1)j'/2$, where $\lceil r \rceil = \min\{k \in \mathbb{Z} : k \geq r\}$ for $r \in \mathbb{R}$. Consequently, the index (i, j) can be computed from k via

$$j = p - \left\lceil \sqrt{p(p+1) - 2k + \frac{9}{4}} - \frac{1}{2} \right\rceil + 1 \quad \text{and} \quad i = k + p - p \cdot j + \frac{j(j-1)}{2}. \tag{9}$$

With the above notation, we can formulate the main result of this paper:

Theorem 1. *We assume that H_0, (1) and (3) hold, and*

$$\int_0^1 E(X_1(t))^4 dt < \infty, \quad \int_0^1 E(X_1^*(t))^4 dt < \infty. \tag{10}$$

Then

$$\frac{NM}{N+M} \widehat{\xi}_{N,M}^T \widehat{L}_{N,M}^{-1} \widehat{\xi}_{N,M} \xrightarrow{\mathcal{D}} \chi^2_{p(p+1)/2}, \quad as \quad N, M \to \infty,$$

where $\chi^2_{p(p+1)/2}$ stands for a χ^2 random variable with $p(p+1)/2$ degrees of freedom.

Theorem 1 implies that the null hypothesis is rejected if the test statistic

$$\widehat{T}_1 = \frac{NM}{N+M} \widehat{\xi}_{N,M}^T \widehat{L}_{N,M}^{-1} \widehat{\xi}_{N,M}$$

exceeds a critical quantile of the chi–square distribution with $p(p+1)/2$ degrees of freedom.
If both samples are Gaussian random processes, the quadratic form $\widehat{\xi}_{N,M}^T \widehat{L}_{N,M}^{-1} \widehat{\xi}_{N,M}$ can be replaced with the normalized sum of the squares of $\widehat{\Delta}_{N,M}(i,j) - \widehat{\Delta}_{N,M}^*(i,j)$, as stated in the following theorem.

Theorem 2. *If X_1, X_1^* are Gaussian processes and the conditions of Theorem 1 are satisfied, then, as $N, M \to \infty$,*

$$\widehat{T}_2 = \frac{NM}{N+M} \sum_{1 \leq i, j \leq p} \frac{1}{2} \frac{\left(\widehat{\Delta}_{N,M}(i,j) - \widehat{\Delta}_{N,M}^*(i,j) \right)^2}{\widehat{\lambda}_i \widehat{\lambda}_j} \xrightarrow{\mathcal{D}} \chi^2_{p(p+1)/2}.$$

5

Observe that the statistic $\widehat{T}_2$ can be written as

$$\widehat{T}_2 = \frac{NM}{N+M}\left\{\sum_{1\le i<j\le p}\frac{\left(\widehat{\Delta}_{N,M}(i,j)-\widehat{\Delta}^*_{N,M}(i,j)\right)^2}{\widehat{\lambda}_i\widehat{\lambda}_j} + \sum_{i=1}^{p}\frac{\left(\widehat{\Delta}_{N,M}(i,i)-\widehat{\Delta}^*_{N,M}(i,i)\right)^2}{2\widehat{\lambda}_i^2}\right\}.$$

Next we discuss the asymptotic consistency of the testing procedure based on Theorem 1. Analogously to the definition of $\widehat{\boldsymbol{\xi}}_{N,M}$ we define the vector $\boldsymbol{\xi} = (\xi(1),\ldots,\xi(p(p+1)/2))$ using the columns of the matrix

$$\boldsymbol{D} = \left(\int_0^1\int_0^1 (C(t,s)-C^*(t,s))\varphi_i(t)\varphi_j(s)dt\,ds\right)_{i,j=1,\ldots,p} \tag{11}$$

instead of $\widehat{\Delta}_N - \widehat{\Delta}^*_M$, i.e.

$$\widehat{\boldsymbol{\xi}} = \mathrm{vech}(\boldsymbol{D}).$$

Theorem 3. *We assume that H_A, (1), (3) and (10) hold. Then there exist random variables $\widehat{h}_1 = \widehat{h}_1(N,M),\ldots,\widehat{h}_{p(p+1)/2} = \widehat{h}_{p(p+1)/2}(N,M)$, taking values in $\{-1,1\}$ such that, as $N,M\to\infty$,*

$$\max_{1\le i\le p(p+1)/2}\left|\widehat{\xi}_{N,M}(i)-\widehat{h}_i\xi(i)\right| = o_P(1) \tag{12}$$

and therefore

$$\left|\widehat{\boldsymbol{\xi}}_{N,M}\right| \xrightarrow{P} |\boldsymbol{\xi}|, \tag{13}$$

where $|\cdot|$ denotes the Euclidean norm. If $\boldsymbol{\xi} \ne \boldsymbol{0}$ and the p largest eigenvalues of C and C^* are positive, we also have

$$\widehat{T}_1 \xrightarrow{P} \infty, \quad as \quad N,M\to\infty. \tag{14}$$

The assumption that the p largest eigenvalues of C and C^* are positive implies that the random functions X_i, $i = 1,\ldots,N$, and X^*_j, $j = 1,\ldots,M$, are not included in a $(p-1)$-dimensional subspace.

The application of the test requires the selection of the number p of the empirical FPC's to be used. A rule of thumb is to choose p so that the first p empirical FPC's in each sample (i.e. those calculated as the eigenfunctions of $\widehat{C}_N$ and $\widehat{C}^*_M$) explain about 85–90% of the variance in each sample. Choosing p too large generally negatively affects the finite sample performance of tests of this type, and for this reason we do not study asymptotics as p tends to infinity. It is often illustrative to apply the test for a range of the values of p; each p specifies a level of relevance of differences in the curves or kernels. A good practical approach is to look at the Karhunen–Loève approximations of the curves in both samples, and choose p which gives approximation errors that can be considered unimportant.

Table 1: Empirical sizes of the tests based on statistics $\widehat{T}_1$ and $\widehat{T}_2$ for non–Gaussian data. The curves in each sample were generated according to (15).

| | $p = 2$ | | | | | |
| | $\widehat{T}_1$ | | | $\widehat{T}_2$ | | |
Sample Sizes	1%	5%	10%	1%	5%	10%
$N = M = 100$	0.005	0.028	0.061	0.152	0.275	0.380
$N = M = 200$	0.003	0.021	0.058	0.163	0.314	0.402
$N = M = 1000$	0.002	0.021	0.056	0.190	0.313	0.426
	$p = 3$					
	$\widehat{T}_1$			$\widehat{T}_2$		
Sample Sizes	1%	5%	10%	1%	5%	10%
$N = M = 100$	0.004	0.028	0.065	0.167	0.332	0.434
$N = M = 200$	0.004	0.024	0.064	0.194	0.338	0.423
$N = M = 1000$	0.004	0.028	0.070	0.240	0.384	0.484

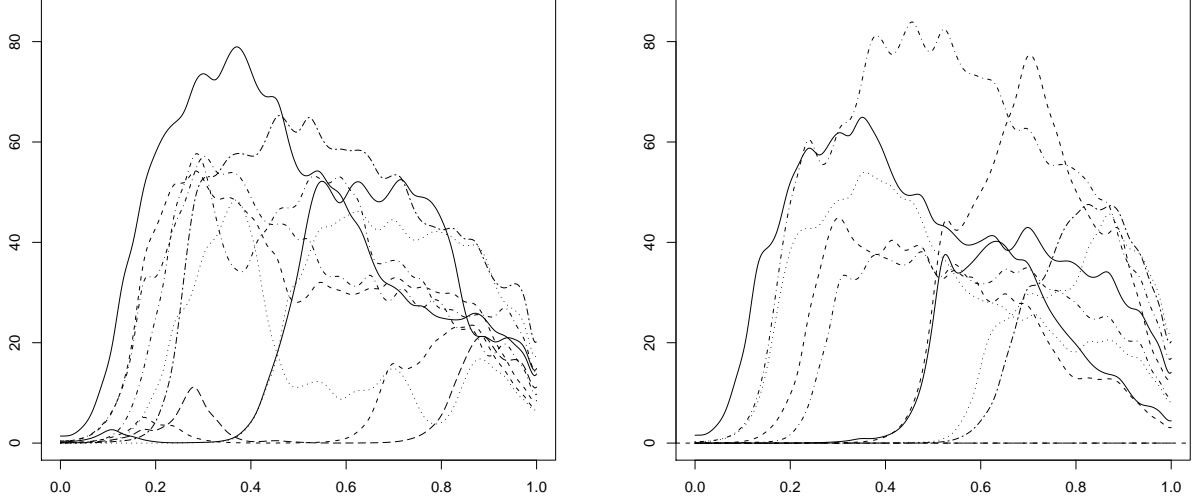## 4. A simulation study and application to medfly data

We first describe the results of a simulations study designed to evaluate finite sample properties of the tests based on the statistics $\widehat{T}_1$ and $\widehat{T}_2$. The emphasis is on the robustness to the violation of the assumption of normality. We simulated Gaussian curves as Brownian motions and Brownian bridges, and non–Gaussian curves via

$$Z_n(t) = A_n \sin(\pi t) + B_n \sin(2\pi t) + C_n \sin(4\pi t), \tag{15}$$

where $A_n = 5Y_1$, $B_n = 3Y_2$, $C_n = Y_3$ and Y_1, Y_2, Y_3 are independent t_5-distributed random variables. All curves were simulated at 1000 equidistant points in the interval $[0, 1]$, and transformed into functional data objects using the Fourier basis with 49 basis functions. For each data generating process we used one thousand replications.

Table 1 displays the empirical sizes for non–Gaussian data. The test based on $\widehat{T}_2$ has severely inflated size, due to the violation of the assumption of normality. As documented in Panaretos et al. (2010), and confirmed by our own simulations, this test has very good empirical size when the data are Gaussian. The test based on $\widehat{T}_1$ is conservative, especially for smaller sample sizes. This is true for both Gaussian and non–Gaussian data; there is not much difference in the empirical size of this test for different data generating processes. Reflecting its conservative size, statistic $\widehat{T}_1$ leads to smaller power than $\widehat{T}_2$. We also studied a Monte Carlo version of the test based on the statistic $\widehat{T}_3 = NM(N + M)^{-1}\widehat{\boldsymbol{\xi}}_{N,M}^{T}\widehat{\boldsymbol{\xi}}_{N,M}$, and found that its finite sample properties were similar to those of

Figure 1: Ten randomly selected smoothed egg-laying curves of short-lived medflies (left panel), and ten such curves for long–lived medflies (right panel).

the test based on $\widehat{T}_1$.

We now describe the results of the application of both tests to an interesting data set consisting of egg–laying trajectories of Mediterranean fruit flies (medflies). The data were kindly made available to us by Hans–Georg Müller. This data set has been extensively studied in biological and statistical literature, see Müller & Stadtmüller (2005) and references therein. We consider 534 egg-laying curves of medflies who lived at least 34 days. We examined two versions of these egg-laying curves, the functions in either version are defined over an interval $[0, 30]$, and $t \leq 30$ is the day. Version 1 curves (denoted $X_i(t)$) are the absolute counts of eggs laid by fly i on day t. Version 2 curves (denoted $Y_i(t)$) are the counts of eggs laid by fly i on day t *relative* to the total number of eggs laid in the lifetime of fly i. The 534 flies are classified into long-lived, i.e. those who lived 44 days or longer, and short-lived, i.e. those who died before the end of the 43rd day after birth. In the data set, there are 256 short-lived, and 278 long-lived flies. This classification naturally defines two samples: *Sample 1:* the egg-laying curves $\{X_i(t)(\text{resp. } Y_i(t)), \ 0 < t \leq 30, \ i = 1, 2, \ldots, 256\}$ of the short-lived flies. *Sample 2:* the egg-laying curves $\{X_j^*(t)(\text{resp. } Y_j^*(t)), \ 0 < t \leq 30, \ j = 1, 2, \ldots, 278\}$ of the long-lived flies. The egg-laying curves are very irregular; Figure 1 shows ten (smoothed) curves of short- and long-lived flies for version 1, Figure 2 shows ten (smoothed) curves for version 2 (both using a B-spline basis for the representation).

Table 2 shows the P–values for the absolute egg-laying counts (version 1). For the statistic $\widehat{T}_1$ the null hypothesis cannot be rejected irrespective of the choice of p. For the statistic $\widehat{T}_2$, the result of the test varies depending on the choice of p. As explained in Section 3., the usual recommendation

8

Figure 2: Ten randomly selected smoothed egg-laying curves of short-lived medflies (left panel), and ten such curves for long–lived medflies (right panel), relative to the number of eggs laid in the fly's lifetime.
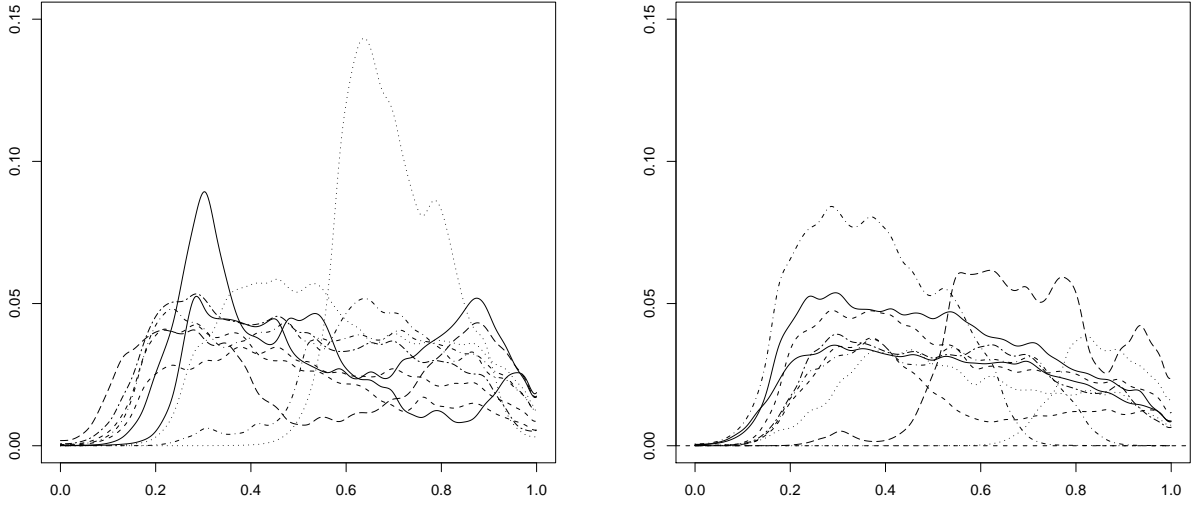

Table 2: P–values (in percent) of the test based on statistics $\widehat{T}_1$ and $\widehat{T}_2$ applied to absolute medfly data. Here f_p denotes the fraction of the sample variance explained by the first p FPCs, i.e. $f_p = (\sum_{k=1}^{p} \hat{\lambda}_k)/(\sum_{k=1}^{N+M} \hat{\lambda}_k)$.

P-values								
p	2	3	4	5	6	7	8	9
$\widehat{T}_1$	82.70	36.22	30.59	63.84	37.71	39.03	33.77	34.77
$\widehat{T}_2$	0.54	0.13	0.11	0.12	0.02	0.00	0.00	0.00
f_p	72.93	78.36	81.87	83.94	85.62	87.08	88.49	89.72

Figure 3: Normal QQ–plots for the scores of the version 2 medfly data with respect to the first two Fourier basis functions. Left – sample 1, Right – sample 2.
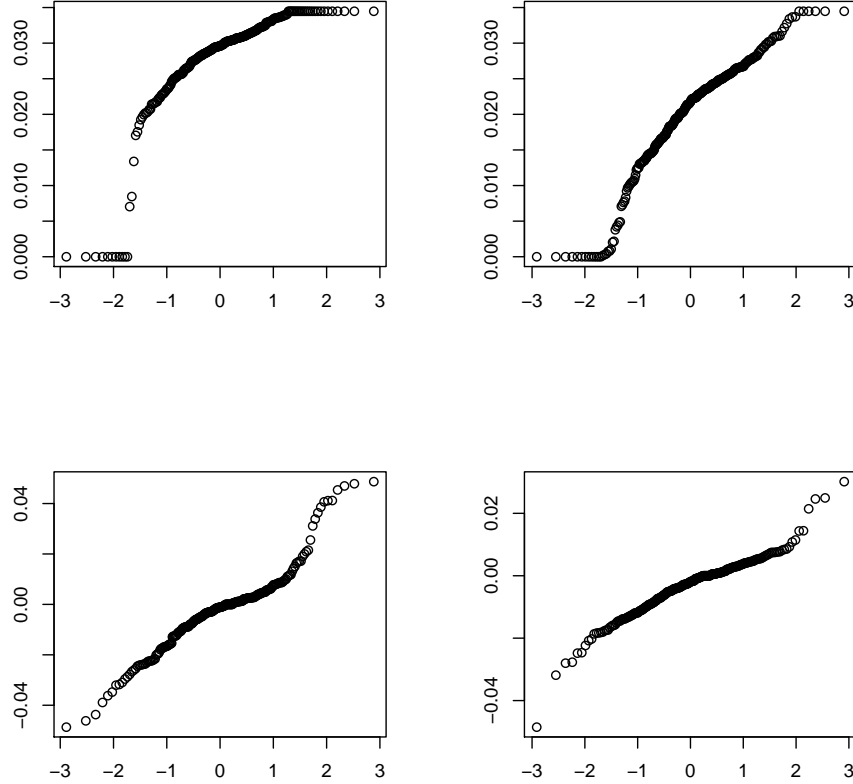

Table 3: P–values (in percent) of the test based on statistics $\widehat{T}_1$ applied to relative medfly data; f_p denotes the fraction of the sample variance explained by the first p FPCs, i.e. $f_p = (\sum_{k=1}^{p} \hat{\lambda}_k)/(\sum_{k=1}^{N+M} \hat{\lambda}_k)$.

P-values							
p	2	3	4	5	6	7	8
$\widehat{T}_1$	0.14	0.06	0.33	1.50	3.79	4.53	10.28
f_p	33.99	44.08	52.72	59.04	65.08	70.40	75.29
p	9	10	11	12	13	14	15
$\widehat{T}_1$	5.51	2.78	5.32	3.21	1.78	6.28	3.80
f_p	79.91	83.72	86.58	89.02	91.34	93.30	95.03

is to use the values of p which explain 85 to 90 percent of the variance. For such values of p, $\widehat{T}_2$ leads to a clear rejection. Since this test has however overinflated size, we conclude that there is little evidence that the covariance structures of version 1 curves for long– and short–lived flies are different. For the version 2 curves, the statistic $\widehat{T}_2$ yields P–values equal to zero (in machine precision), potentially indicating that the covariance structures for the short– and long–lived flies are different. The assumption of a normal distribution is however questionable, as the QQ-plots in Figure 3 show. These QQ-plots are constructed for the inner products $\langle Y_i, e_k \rangle$ and $\langle Y_i^*, e_k \rangle$, where the Y_i are the curves from one of the samples (we cannot pool the data to construct QQ-plots because we test if the stochastic structures are different), and e_k is the kth element of the Fourier basis. The normality of a functional sample implies the normality of all projections onto a complete orthonormal system . For $\langle X_i, e_k \rangle$, the QQ-plots show a strong deviation from a straight line for some projections. Almost all projections $\langle Y_i, e_k \rangle$ have QQ-plots indicating a strong deviation from normality. It is therefore important to apply the robust test based on the statistic $\widehat{T}_1$. The corresponding P–values for version 2 are displayed in Table 3. For most values of p, these P–values indicate the rejection of H_0. Many of them hover around the 5 percent level, but since the test is conservative, we can with confidence view them as favoring H_A.

The above application confirms the properties of the statistics established through the simulation study. It shows that while there is little evidence that the covariance structures for the absolute counts are different, there is strong evidence that they are different for relative counts.

5.   Proofs of the results of Section 3.

The proof of Theorem 1 follows from several lemmas, which we establish first. We can and will assume without loss of generality that $\mu(t) = \mu^*(t) = 0$ for all $t \in [0, 1]$.

We will use the identity

$$
\frac{1}{N^{1/2}} \sum_{k=1}^{N} \left(X_k(t) - \overline{X}_N(t) \right) \left(X_k(s) - \overline{X}_N(s) \right) \tag{16}
$$
$$
= \frac{1}{N^{1/2}} \sum_{k=1}^{N} X_k(t) X_k(s) - N^{1/2} \overline{X}_N(t) \overline{X}_N(s),
$$

and an analogous identity for the second sample.

Our first lemma establishes bounds in probability which will often be used in the proofs.

Lemma 1. *Under the assumptions of Theorem 1, as $N, M \to \infty$,*

$$
\left\| N^{-1/2} \sum_{k=1}^{N} \{ X_k(t) X_k(s) - C(t, s) \} \right\| = O_P(1), \tag{17}
$$
$$
\left\| N^{1/2} \overline{X}_N(t) \right\| = O_P(1), \tag{18}
$$

and

$$\left\| M^{-1/2} \sum_{k=1}^{M} \{X_k^*(t)X_k^*(s) - C^*(t,s)\} \right\| = O_P(1), \tag{19}$$

$$\left\| M^{1/2}\overline{X}_M^*(t) \right\| = O_P(1). \tag{20}$$

Proof. First we note that

$$E \int_0^1 \int_0^1 \left[\frac{1}{N^{\frac{1}{2}}} \sum_{k=1}^{N} \{X_k(t)X_k(s) - C(t,s)\} \right]^2 dt\,ds = \int_0^1 \int_0^1 E\{X_1(t)X_1(s) - C(t,s)\}^2 dt\,ds,$$

so, by Markov's inequality, we have

$$\left\| \frac{1}{N^{\frac{1}{2}}} \sum_{k=1}^{N} \{X_k(t)X_k(s) - C(t,s)\} \right\|^2 = O_P(1).$$

Similar arguments yield $(18) - (20)$. $\qquad\square$

The next lemma shows that the estimation of the mean functions, cf. the definition of the projections $\widehat{a}_k(i)$ and $\widehat{a}_k^*(j)$ in (5) and (6), has an asymptotically negligible effect.

Lemma 2. *Under the assumptions of Theorem 1, for all $1 \le i, j \le p$, as $N, M \to \infty$,*

$$N^{1/2}\widehat{\Delta}_N(i,j) = \frac{1}{N^{1/2}} \sum_{k=1}^{N} \langle X_k, \widehat{\varphi}_i \rangle \langle X_k, \widehat{\varphi}_j \rangle + O_P\left(N^{-1/2}\right)$$

and

$$M^{1/2}\widehat{\Delta}_M^*(i,j) = \frac{1}{M^{1/2}} \sum_{k=1}^{M} \langle X_k^*, \widehat{\varphi}_i \rangle \langle X_k^*, \widehat{\varphi}_j \rangle + O_P\left(M^{-1/2}\right).$$

Proof. Using (16) and (18) we have by the Cauchy-Schwarz inequality,

$$\left| \int_0^1 \int_0^1 N^{1/2}\overline{X}_N(t)\overline{X}_N(s)\widehat{\varphi}_i(t)\widehat{\varphi}_j(s)dt\,ds \right|$$

$$= N^{-1/2}\left| \int_0^1 N^{1/2}\overline{X}_N(t)\widehat{\varphi}_i(t)dt \right| \left| \int_0^1 N^{1/2}\overline{X}_N(s)\widehat{\varphi}_j(s)ds \right|$$

$$\le N^{-1/2}\left(\int_0^1 \left(N^{1/2}\overline{X}_N(t)\right)^2 dt \int_0^1 \widehat{\varphi}_i^2(t)dt \right)^{1/2} \left(\int_0^1 \left(N^{1/2}\overline{X}_N(s)\right)^2 ds \int_0^1 \widehat{\varphi}_j^2(s)ds \right)^{1/2}$$

$$= N^{-1/2} \int_0^1 \left(N^{1/2}\overline{X}_N(t)\right)^2 dt$$

$$= O_P\left(N^{-1/2}\right).$$

The second part can be proven in the same way. $\qquad\square$

We now state bounds on the distances between the estimated and the population eigenvalues and eigenfunctions. These bounds are true under the null hypothesis, and extend the corresponding one sample bounds.

Lemma 3. *If the conditions of Theorem 1 are satisfied, then, as $N, M \to \infty$,*

$$\max_{1 \le i \le p} |\widehat{\lambda}_i - \lambda_i| = O_P\left((N + M)^{-1/2}\right)$$

and

$$\max_{1 \le i \le p} \|\widehat{\varphi}_i - \widehat{c}_i \varphi_i\| = O_P\left((N + M)^{-1/2}\right),$$

where

$$\widehat{c}_i = \widehat{c}_i(N, M) = \text{sign}(\langle \widehat{\varphi}_i, \varphi_i \rangle).$$

Proof. It follows from $(17) - (20)$ and the assumption $C = C^*$ that

$$\left\|\widehat{R}_{N,M} - C\right\| = O_P\left(\left(N^{1/2} + M^{1/2}\right) \big/ (N + M)\right),$$

and since $N^{1/2} + M^{1/2} \le 2(N + M)^{1/2}$, the result follows from the corresponding one sample bounds, see e.g. Chapter 2 of Horváth & Kokoszka (2011+). $\qquad \square$

Lemma 3 now allows us to replace the estimated eigenfunctions by their population counterparts. The random signs $\widehat{c}_i$ must appear in the formulation of Lemma 4, but they cancel in the subsequent results.

Lemma 4. *If the conditions of Theorem 1 are satisfied, then, for all $1 \le i, j \le p$, as $N, M \to \infty$,*

$$\left(\frac{NM}{N + M}\right)^{1/2} \left(\widehat{\Delta}_N(i, j) - \widehat{\Delta}_M^*(i, j)\right)$$

$$= \left(\frac{NM}{N + M}\right)^{1/2} \left\{\frac{1}{N} \sum_{k=1}^{N} \langle X_k, \widehat{c}_i \varphi_i \rangle \langle X_k, \widehat{c}_j \varphi_j \rangle - \frac{1}{M} \sum_{k=1}^{M} \langle X_k^*, \widehat{c}_i \varphi_i \rangle \langle X_k^*, \widehat{c}_j \varphi_j \rangle\right\} + o_P(1).$$

Proof. We write

$$\frac{1}{N} \sum_{k=1}^{N} \langle X_k, \widehat{\varphi}_i \rangle \langle X_k, \widehat{\varphi}_j \rangle - \int_0^1 \int_0^1 C(t, s) \widehat{\varphi}_i(t) \widehat{\varphi}_j(s) dt \, ds$$

$$= N^{1/2} \int_0^1 \int_0^1 \left\{\frac{1}{N^{1/2}} \sum_{k=1}^{N} \left(X_k(t) X_k(s) - C(t, s)\right)\right\} \widehat{\varphi}_i(t) \widehat{\varphi}_j(s) dt \, ds.$$

Using Lemmas $1 - 3$ we get

$$\left|\int_0^1 \int_0^1 \left\{\frac{1}{N^{1/2}} \sum_{k=1}^{N} \left(X_k(t) X_k(s) - C(t, s)\right)\right\} \left(\widehat{\varphi}_i(t) \widehat{\varphi}_j(s) - \widehat{c}_i \varphi_i(t) \widehat{c}_j \varphi_j(s)\right) dt \, ds\right|$$

$$= \left|\int_0^1 \int_0^1 \left\{\frac{1}{N^{1/2}} \sum_{k=1}^{N} \left(X_k(t) X_k(s) - C(t, s)\right)\right\}\right.$$

$$\left. \times \left\{\left(\widehat{\varphi}_i(t) - \widehat{c}_i \varphi_i(t)\right) \widehat{\varphi}_j(s) + \widehat{c}_i \varphi_i(t) \left(\widehat{\varphi}_j(s) - \widehat{c}_j \varphi_j(s)\right)\right\} dt \, ds\right|$$

$$\leq \left(\int_0^1 \int_0^1 \left\{ \frac{1}{N^{1/2}} \sum_{k=1}^N \left(X_k(t) X_k(s) - C(t,s) \right) \right\}^2 dt\, ds \right.$$

$$\times \left. \int_0^1 \int_0^1 \left(\widehat{\varphi}_i(t) - \widehat{c}_i \varphi_i(t) \right)^2 \widehat{\varphi}_j^2(s) dt\, ds \right)^{1/2}$$

$$+ \left(\int_0^1 \int_0^1 \left\{ \frac{1}{N^{1/2}} \sum_{k=1}^N \left(X_k(t) X_k(s) - C(t,s) \right) \right\}^2 dt\, ds \right.$$

$$\times \left. \int_0^1 \int_0^1 \varphi_i^2(t) \left(\widehat{\varphi}_j(s) - \widehat{c}_j \varphi_j(s) \right)^2 dt\, ds \right)^{1/2}$$

$$= \left\| \frac{1}{N^{1/2}} \sum_{k=1}^N \left(X_k(t) X_k(s) - C(t,s) \right) \right\| \left\{ \left\| \widehat{\varphi}_i - \widehat{c}_i \varphi_i \right\| + \left\| \widehat{\varphi}_j - \widehat{c}_j \varphi_j \right\| \right\}$$

$$= o_P(1).$$

Similar arguments give that

$$\left| \int_0^1 \int_0^1 \left\{ \frac{1}{M^{1/2}} \sum_{k=1}^M \left(X_k^*(t) X_k^*(s) - C^*(t,s) \right) \right\} \left\{ \widehat{\varphi}_i(t) \widehat{\varphi}_j(s) - \widehat{c}_i \varphi_i(t) \widehat{c}_j \varphi_j(s) \right\} dt\, ds \right| = o_P(1).$$

Since $C = C^*$, the lemma is proven. $\qquad\square$

The previous lemmas isolated the main terms in the differences $\widehat{\Delta}_N(i,j) - \widehat{\Delta}_M^*(i,j)$. The following lemma describes the limits of these main terms (without the random signs).

Lemma 5. *If the conditions of Theorem 1 are satisfied, then, as $N, M \to \infty$,*

$$\{\Delta_{N,M}(i,j), 1 \leq i, j \leq p\} \xrightarrow{\mathcal{D}} \{\Delta(i,j), 1 \leq i, j \leq p\},$$

where

$$\Delta_{N,M}(i,j) = \left(\frac{NM}{N+M} \right)^{1/2} \left\{ \frac{1}{N} \sum_{k=1}^N \langle X_k, \varphi_i \rangle \langle X_k, \varphi_j \rangle - \frac{1}{M} \sum_{k=1}^M \langle X_k^*, \varphi_i \rangle \langle X_k^*, \varphi_j \rangle \right\},$$

and $\{\Delta(i,j), 1 \leq i, j \leq p\}$ is a Gaussian matrix with $E\Delta(i,j) = 0$ and

$$E\Delta(i,j)\Delta(i',j') = (1 - \Theta) \left\{ E\left(\langle X_1, \varphi_i \rangle \langle X_1, \varphi_j \rangle \langle X_1, \varphi_{i'} \rangle \langle X_1, \varphi_{j'} \rangle \right) \right.$$

$$\left. - E\left(\langle X_1, \varphi_i \rangle \langle X_1, \varphi_j \rangle \right) E\left(\langle X_1, \varphi_{i'} \rangle \langle X_1, \varphi_{j'} \rangle \right) \right\}$$

$$+ \Theta \left\{ E\left(\langle X_1^*, \varphi_i \rangle \langle X_1^*, \varphi_j \rangle \langle X_1^*, \varphi_{i'} \rangle \langle X_1^*, \varphi_{j'} \rangle \right) \right.$$

$$\left. - E\left(\langle X_1^*, \varphi_i \rangle \langle X_1^*, \varphi_j \rangle \right) E\left(\langle X_1^*, \varphi_{i'} \rangle \langle X_1^*, \varphi_{j'} \rangle \right) \right\}.$$

Proof. First we note that

$$E\langle X_1, \varphi_i \rangle \langle X_1, \varphi_j \rangle = E\langle X_1^*, \varphi_{i'} \rangle \langle X_1^*, \varphi_{j'} \rangle = \begin{cases} 0 & \text{if } i \neq j, \\ \lambda_i & \text{if } i = j. \end{cases}$$

Since $E\left(\langle X_1, \varphi_i \rangle \langle X_1, \varphi_j \rangle \right)^2 < \infty$ and $E\left(\langle X_1^*, \varphi_i \rangle \rangle X_1^*, \varphi_j \rangle \right)^2 < \infty$, the multivariate central limit theorem implies the result. $\qquad\square$

Finally, we need an asymptotic approximation to the covariances $\widehat{L}_{N,M}(k, k')$. Let

$$L_{N,M}(k, k') = (1 - \Theta_{N,M}) \left\{ \frac{1}{N} \sum_{\ell=1}^{N} a_\ell(i)a_\ell(j)a_\ell(i')a_\ell(j') - \langle \widehat{\mathfrak{C}}_N \widehat{\varphi}_i, \widehat{\varphi}_j \rangle \langle \widehat{\mathfrak{C}}_N \widehat{\varphi}_{i'}, \widehat{\varphi}_{j'} \rangle \right\}$$

$$+ \Theta_{N,M} \left\{ \frac{1}{M} \sum_{\ell=1}^{M} a_\ell^*(i)a_\ell^*(j)a_\ell^*(i')a_\ell^*(j') - \langle \widehat{\mathfrak{C}}_M^* \widehat{\varphi}_i, \widehat{\varphi}_j \rangle \langle \widehat{\mathfrak{C}}_M^* \widehat{\varphi}_{i'}, \widehat{\varphi}_{j'} \rangle \right\},$$

where

$$a_\ell(i) = \langle X_\ell, \varphi_i \rangle \quad \text{and} \quad a_\ell^*(i) = \langle X_\ell^*, \varphi_i \rangle,$$

and i, j, i', j' are determined from k and k' as in (8) and (9).

Lemma 6. *If the conditions of Theorem 1 are satisfied, then for all $1 \le k, k' \le p(p+1)/2$,*

$$\widehat{L}_{N,M}(k, k') - \widehat{c}_i \widehat{c}_j \widehat{c}_{i'} \widehat{c}_{j'} L_{N,M}(k, k') = o_P(1) \quad as \quad N, M \to \infty,$$

where (i, j) and (i', j') are determined from k and k' as in (8) and (9).

Proof. The result follows from Lemma 3 along the lines of the proof of Lemma 4 $\qquad\square$

Proof of Theorem 1. According to Lemma 2 and Lemmas 4 – 6, the asymptotic distribution of $\widehat{\boldsymbol{\xi}}_{N.M}^T \widehat{L}_{N,M}^{-1} \widehat{\boldsymbol{\xi}}_{N,M}$ does not depend on the signs $\widehat{c}_1, \ldots, \widehat{c}_p$, so it is sufficient to prove the result for $\widehat{c}_1 = \ldots = \widehat{c}_p = 1$. The law of large numbers yields that

$$L_{N,M}(k, k') \xrightarrow{P} L(k, k'), \tag{21}$$

where

$$L(k, k') = (1 - \Theta) \left\{ E\big(a_1(i)a_1(j)a_1(i')a_1(j')\big) - E\big(a_1(i)a_1(j)a_1(i')a_1(j')\big) \right\} \tag{22}$$

$$+ \Theta \left\{ E\big(a_1^*(i)a_1^*(j)a_1^*(i')a_1^*(j')\big) - E\big(a_1^*(i)a_1^*(j)a_1^*(i')a_1^*(j')\big) \right\}.$$

Now the result follows from Lemmas 2, 4 and 5 $\qquad\square$

Proof of Theorem 2. We continue to assume that $\mu = \mu^* = 0$. This means that, under H_0, $X_1, \ldots, X_N, X_1^*, \ldots, X_M^*$ are independent and identically distributed Gaussian processes. Hence $\langle X_1, \varphi_i \rangle = \lambda_i^{1/2} N_i$, $\langle X_1, \varphi_j \rangle = \lambda_j^{1/2} N_j^*$, where $N_i, N_j^*, 1 \le i, j \le p$ are independent standard normal random variables. We have already pointed out that

$$E\langle X_1, \varphi_i \rangle \langle X_1, \varphi_j \rangle = \begin{cases} 0 & \text{if } i \ne j, \\ \lambda_i & \text{if } i = j, \end{cases}$$

and

$$E\langle X_1, \varphi_{i'} \rangle \langle X_1, \varphi_{j'} \rangle = \begin{cases} 0 & \text{if } i' \ne j', \\ \lambda_{i'} & \text{if } i = j'. \end{cases}$$

15

If $i = j = i' = j'$, then

$$E\big(\langle X_1, \varphi_i\rangle\big)^4 - \big(E(\langle X_1, \varphi_i\rangle)^2\big)^2 = \lambda_i^2 E\big(N_i^4 - (EN_i^2)^2\big) = 2\lambda_i^2.$$

If $i = i'$ and $j = j'$ $(i \neq j)$, then

$$E\langle X_1, \varphi_i\rangle\langle X_1, \varphi_j\rangle\langle X_1, \varphi_{i'}\rangle\langle X_1, \varphi_{j'}\rangle = \lambda_i\lambda_j.$$

In all other cases

$$E\langle X_1, \varphi_i\rangle\langle X_1, \varphi_j\rangle\langle X_1, \varphi_{i'}\rangle\langle X_1, \varphi_{j'}\rangle = 0.$$

Hence $\Delta(i,j)$, $1 \leq i \leq j \leq p$, are independent normal random variables with mean 0 and

$$E\Delta^2(i,j) = \begin{cases} \lambda_i\lambda_j & \text{if} \quad i \neq j, \\ 2\lambda_i^2 & \text{if} \quad i = j. \end{cases}$$

Now the result follows from Lemmas $1 - 5$. $\qquad\square$

Proof of Theorem 3. First we observe that by the law of large numbers we have

$$\int_0^1 \int_0^1 (\widehat{R}_{N,M}(t,s) - R(t,s))^2 dt\, ds = o_P(1).$$

Hence using the result in section VI.1. of Gohberg et al. (1990) (cf. Lemmas 2.2 and 2.3 in Horváth & Kokoszka (2011+)) we get that

$$\max_{1 \leq i \leq p}\left|\widehat{\lambda}_i - \lambda_i\right| = o_P(1) \tag{23}$$

and

$$\max_{1 \leq i \leq p}\|\widehat{\varphi}_i - \widehat{c}_i\varphi_i\| = o_P(1), \tag{24}$$

where $\widehat{c}_i = \widehat{c}_i(N,M) = \text{sign}(\langle\widehat{\varphi}_i, \varphi_i\rangle)$. Relations (23) and (24) show that Lemma 3 remains true. It follows from the law of large numbers and (24) that for all $1 \leq i, j \leq p$

$$\left|\widehat{\Delta}_N(i,j) - \widehat{\Delta}_M^*(i,j) - \widehat{c}_i\widehat{c}_j \int_0^1 \int_0^1 (C(t,s) - C^*(t,s))\, \varphi_i(t)\varphi_j(s)dt\, ds\right|$$

$$= \left|\int_0^1 \int_0^1 \left(\widehat{C}_N(t,s) - \widehat{C}_M^*(t,s)\right)\widehat{\varphi}_i(t)\widehat{\varphi}_j(s)dt\, ds - \widehat{c}_i\widehat{c}_j \int_0^1 \int_0^1 (C(t,s) - C^*(t,s))\, \varphi_i(t)\varphi_j(s)dt\, ds\right|$$

$$\leq \left|\int_0^1 \int_0^1 \left(\widehat{C}_N(t,s) - C(t,s) - \left(\widehat{C}_M^*(t,s) - C^*(t,s)\right)\right)\widehat{\varphi}_i(t)\widehat{\varphi}_j(s)dt\, ds\right|$$

$$+ \left|\int_0^1 \int_0^1 (C(t,s) - C^*(t,s))\, (\widehat{\varphi}_i(t)\widehat{\varphi}_j(s) - \widehat{c}_i\varphi_i(t)\widehat{c}_j\varphi_j(s))\, dt\, ds\right|$$

16

$$\leq \left\| \widehat{C}_N - C \right\| + \left\| \widehat{C}_M^* - C^* \right\| + \| C - C^* \| \, \| \widehat{\varphi}_i \widehat{\varphi}_j - \widehat{c}_i \varphi_i \widehat{c}_i \varphi_j \|$$
$$= o_P(1),$$

where the fact that $\|\varphi_i\| = 1 = \|\widehat{\varphi}_i\|$ was used. Hence the proof of (12) is complete. It is also clear that (12) implies (13).

Next we observe that Lemma 6 and (21) remain true under the alternative. Now by some lengthy calculations it can be verified that L given in (22) is positive definite so that (14) follows from (13). $\square$

References

Abadir, K. M. & Magnus, J.R. (2005). *Matrix algebra.* Cambridge University Press, New York.

Benko, M., Härdle, W. & A. Kneip (2009). Common functional principal components. *Ann. Statist.*, **37**, 1-34.

Bosq, D. (2000). *Linear processes in function spaces.* Springer, New York.

Ferraty, F. & Romain, Y., editors, (2011). *The Oxford handbook of functional data analysis.* Oxford University Press.

Ferraty, F. & Vieu, P. (2006). *Nonparametric functional data analysis: Theory and practice.* Springer, New York.

Gabrys, R., Horváth, L. & Kokoszka, P. (2010). Tests for error correlation in the functional linear model. *J. Amer. Statist. Assoc.*, **105**, 1113-1125.

Gervini, D. (2008). Robust functional estimation using the spatial median and spherical principal components. *Biometrika*, **95**, 587-600.

Gohberg, I., Goldberg, S. & Kaashoek, M.A. (1990). *Classes of linear operators.* Operator Theory: Advances and Applications, **49**, Birkhäuser, Basel.

Horváth, L. & Kokoszka, P. (2011+). *Inference for functional data with applications.* Springer Series in Statistics. Springer, New York. Forthcoming.

Horváth, L., Kokoszka, P. & Reeder, R. (2011). Estimation of the mean of functional time series and a two sample problem. Technical report, University of Utah.

Horváth, L., Kokoszka, P. & Reimherr, M. (2009). Two sample inference in functional linear models. *Canad. J. Statist.*, **37**, 571-591.

Müller, H-G. & Stadtmüller, U. (2005). Generalized functional linear models. *Ann. Statist.*, **33**, 774-805.

Panaretos, V. M., Kraus, D. & Maddocks, J. H. (2010). Second-order comparison of Gaussian random functions and the geometry of DNA minicircles. *J. Amer. Statist. Assoc.*, **105**, 670–682.

Ramsay, J., Hooker, G. & Graves, S. (2009). *Functional data analysis with R and MATLAB*. Springer, New York.

Ramsay, J. O. & Silverman, B. W. (2005). *Functional data analysis*. Springer, New York.

Reiss, P. T. & Ogden, R. T. (2007). Functional principal component regression and functional partial least squares. *J. Amer. Statist. Assoc.*, **102**, 984-996.

Yao, F. & Müller, H-G. (2010). Functional quadratic regression. *Biometrika*, **97**, 49-64.